A Spatial Modeling Framework for Monitoring Surveys with Different Sampling Protocols with a Case Study for Bird Abundance in Mid-Scandinavia

Jorge SICACHA-PARADA, Diego PAVON-JORDAN, Ingelin STEINSLAND, Roel MAY, Bård STOKKE, and Ingar Jostein ØIEN

Quantifying the total number of individuals (abundance) of species is the basis for spatial ecology and biodiversity conservation. Abundance data are mostly collected through professional surveys as part of monitoring programs, often at a national level. These surveys rarely follow exactly the same sampling protocol in different countries, which represents a challenge for producing biogeographical abundance maps based on the transboundary information available covering more than one country. Moreover, not all species are properly covered by a single monitoring scheme, and countries typically collect abundance data for target species through different monitoring schemes. We present a new methodology to model total abundance by merging count data information from surveys with different sampling protocols. The proposed methods are used for data from national breeding bird monitoring programs in Norway and Sweden. Each census collects abundance data following two different sampling protocols in each country, i.e., these protocols provide data from four different sampling processes. The modeling framework assumes a common Gaussian Random Field shared by both the observed and true abundance with either a linear or a relaxed linear association between them. The models account for particularities of each sampling protocol by including terms that affect each observation process, i.e., accounting for differences in observation units and detectability. Bayesian inference is performed using the Integrated Nested Laplace Approximation (INLA) and the Stochastic Partial Differential Equation (SPDE) approach for spatial modeling. We also present the results of a simulation study based on the empirical census data from mid-Scandinavia to assess the performance of the models under model misspecification. Finally, maps of the expected abundance of birds in our study region in

J. Sicacha-Parada (✉) · I. Steinsland, Department of Mathematical Sciences, Norwegian University of Science and Technology (NTNU), Trondheim, Norway
(E-mail: *jorge.sicacha@ntnu.no; ingelin.steinsland@ntnu.no*)

D. Pavon-Jordan · R. May · B. Stokke, Department of Terrestrial Ecology, Norwegian Institute for Nature Research (NINA), P.O. Box 5685, Torgarden, 7485 Trondheim, Norway

I. J. Øien
Norwegian Ornithological Society-BirdLife Norway, Sandgata 30 B, 7012 Trondheim, Norway.

mid-Scandinavia are presented with uncertainty estimates. We found that the framework allows for consistent integration of data from surveys with different sampling protocols. Further, the simulation study showed that models with a relaxed linear specification are less sensitive to misspecification, compared to the model that assumes linear association between counts. Relaxed linear specifications of total bird abundance in mid-Scandinavia improved both goodness of fit and the predictive performance of the models.

Supplementary materials accompanying this paper appear on-line.

Key Words: Data integration; Joint species distribution models; Bayesian statistics; Latent Gaussian Models; Gaussian Random Fields.

1. INTRODUCTION

Understanding why organisms are where they are and what drives changes in their abundances is one of the main pillars of spatial ecology (Brodie et al. 2020) and is critical to propose effective measures to preserve biodiversity. In this regard, species distribution models (SDMs) have typically been used to gain a better understanding of species–habitat relationships (Brodie et al. 2020; Bradter et al. 2021) and to guide conservation practitioners and policy makers (Araujo et al. 2019). Previous SDMs using abundance data have revealed higher predictive performance in comparison with those using occurrence data (Howard et al. 2014; Johnston et al. 2015). Yet, the majority of SDMs published to date used presence/absence (i.e., occurrence) data (Araujo et al. 2019; Yu et al. 2020), rather than abundance data (count of individuals), especially in large-scale studies (Miller et al. 2019). This limits our ability to robustly infer, for example, regions with high density of individuals (Johnston et al. 2015), which is of paramount importance in conservation (Massimino et al. 2017). For example, estimating abundance hotspots can inform and help authorities to select sites that may qualify to be included in the network of protected areas. Indeed, one of the main criteria to identify important areas for conservation under the European Union's Bird Directive (i.e., Special Protection Areas; SPA) is that a site accommodates regularly 1% of the total biogeographical population of a species of conservation concern or more than 20,000 individuals of wetland birds (EU's Birds Directive, 2009/147/EC 2009). Moreover, this Directive states that *"The measures to be taken must apply to various factors which may affect the numbers of birds, namely the repercussions of man's activities and in particular the destruction and pollution of their habitats[...]"*. Abundance data can also be useful to detect and predict areas where human-wildlife conflicts may arise (e.g., May et al. 2020), informing the corresponding authorities that infrastructure and further human development such as siting of powerlines and wind farms must be planned carefully (e.g., De Lucas et al. 2008; May et al. 2020). Information about abundance is ultimately requested by national (e.g., Directorates, Environmental Agencies) and international (e.g. European Commission) authorities as basis to propose biodiversity conservation policies at different scales. This information should be based on all available count data.

Most countries have monitoring programs following national law and as signatories of international biodiversity conservation Directives and Conventions. These different national monitoring schemes may cover the same taxon (e.g., most countries have a national monitoring scheme for breeding birds) but can differ in the species recorded (different set

of species may occur in different countries and at different densities) and, most importantly, they usually follow different sampling protocols, which makes the information obtained by different schemes not directly comparable. Furthermore, not all species are well represented in the data gathered within a single 'general' protocol. For this reason, many countries have, for example, additional targeted monitoring schemes that complement the information for species that are considered poorly represented in the more general monitoring scheme, for example, colonial birds such as herons in Greece, raptors and waterbirds in Finland, nocturnal birds in Spain; see also Buckland and Johnston (2017). National common bird monitoring schemes and those targeting particular (groups of) species provide together the largest datasets known on species abundance in time and space. However, at the (sub)national level, these datasets have mainly been used independently (Kålås 2010; Bevanger et al. 2014; Kéry and Royle 2009; Soykan et al. 2016) and multi-country studies have mostly analyzed these data either independently for each country to later draw common conclusions from the country-specific estimates (Lehikoinen et al. 2019) or combining the raw data with limited account for sampling differences (e.g., total abundance of waders; Lindström et al. 2019). Thus, overlooking the potential of integrating such a large amount of standardized data seems like an under usage of the effort and resources spent in collecting these data, especially when the taxa included in such monitoring schemes are very diverse, allowing not only to carry out species-specific analyses but also, potentially, community-level studies. This study was motivated by the need for estimates of the total abundance of birds in mid-Scandinavia based on high quality (i.e., standardized surveys) localized data on bird abundances from the common breeding bird monitoring programs in Norway (TOV-E) and Sweden (BBS). An estimate of the total abundance of birds can be used as an input for models that inform on the risk of infrastructure development (e.g., new powerlines and wind farms) for birds. The TOV-E and the BBS both provide standardized count data, but they differ in their sampling protocols. Both countries collect observations in point counts and transect surveys. In Norway, the main point counts (all species recorded) are complemented with line transects (only a subset of 'rare' species also included in point counts are recorded—see further details in Sect. 2). However, in Sweden, the line transects and the point counts can be regarded as two different censuses (i.e., all species are counted in both census methods). These differences present the challenge of integrating the four sources of spatial information (points and transects in both Norway and Sweden) with different sampling protocols into one estimate for the spatial distribution of bird abundance for the entire region of interest (Brodie et al. 2020; Gruss and Thorson 2019).

The scarcity of studies applying large-scale abundance SDMs is likely related to (i) the generally lower availability of abundance data compared to occurrence data for most species (Miller et al. 2019; Buckland and Johnston 2017 and references therein) and (ii) statistical and computational challenges of modeling abundance data. Great methodological advancements to overcome some of these problems have been developed in the past decade, especially for integrating different data types, see Miller et al. (2019) and references therein. Most of these efforts have focused on enabling the use of casually collected (non-standardized) presence-only data to increase spatial coverage and data points of certain species (see also Buckland and Johnston 2017). The possibility of improving SDMs by integrating abundance (count) data collected under different standardized monitoring schemes

is most often neglected. Thus, in addition to the integration of data from different countries, merging data from different schemes (from one or several countries) can thus improve the estimates of abundance obtained from all available count data.

Given the existing gap in methodology for proper integration of standardized count data, we here propose a generic modeling framework that integrates standardized count data from various monitoring schemes (i.e., designed surveys) with different sampling protocols. The models can ultimately produce one single estimate of abundance (total abundance of birds in our case study) and its uncertainty based in data from different sampling protocols. In addition, it also gives interpretable estimates of the ecological parameters driving this abundance. Our methodology, thus, analyzes these data in a unique, single framework to produce models that account for different sampling processes, and describe and predict the spatial distribution of abundance.

Spatial modeling of multiple data sources has been approached for example in the context of coregionalization models (Banerjee et al. 2015; Blangiardo and Cameletti 2015; Krainski 2019) and recently reviewed in Miller et al. (2019). These are multivariate models for measurements that vary jointly over a region and have been defined through a hierarchical structure and fitted using Markov Chain Monte Carlo (MCMC) techniques (Banerjee et al. 2015). For the family of Spatial Latent Gaussian Models (Rue and Held 2005), the INLA-SPDE approach (Rue et al. 2009; Lindgren et al. 2011) and its easy implementation in the INLA library of R have emerged as a faster alternative to jointly model multiple sources of information. Such method has been applied to multivariate models related with, for example, air pollution data (Cameletti et al. 2019), and hydrology (Roksvåg et al. 2020). The proposed framework framework assumes the existence of a latent process, underlying all the observed abundances, that represents the true expected abundances. The true expected abundance varies in space through spatial covariates as well as a spatial random effect. Given the true expected abundance, we assume that the observed abundances follow Poisson distributions. For each observation process a linear relation between the expected counts and the true expected abundances is assumed. Further, we assume the existence of a common spatial random effect that drives the observed counts (cf. Miller et al. 2019) for all the observation processes. Given that the linear assumption may not depict the true relationship between the expected counts and the true expected abundances, we also propose models that allow deviations from this assumption. The proposed models are suitable doing computational fast inference using the INLA-SPDE approach, which approximates the posterior densities of parameters and hyperparameters.

To the best of our knowledge, methodologies for jointly modeling spatial abundance using data from multi-country standardized biodiversity monitoring programs with different sampling protocols have not been published before. By properly integrating data from different monitoring schemes, our method can be part of solving some of the issues inherent to monitoring data raised in Buckland and Johnston (2017), such as the scarcity of data, low representability, and small geographical scale. This opens new possibilities for more robust international assessments of species distributions and abundance using count data from diverse national monitoring programs, which is of paramount importance for understanding global change impacts on biodiversity (Buckland and Johnston 2017; Massimino et al. 2017). We validate this framework with a case study aiming at estimating total bird

abundance in mid-Scandinavia and a simulation study that explores the effects of misspec-
ification on the proposed models.

This paper is organized as follows: In Sect. 2, we describe the data from the Norwegian and
Swedish monitoring programs in detail. Moreover, we explain how we preprocessed these
census data, present an exploratory analysis and introduce the set of candidate explanatory
variables for our models. In Sect. 3, models as well as inference methodology and measures
for evaluating and comparing models are presented. In Sect. 4, we set up a simulation study
to explore how the proposed models perform in scenarios with different relation between
the observed and the true abundances. In Sect. 5, results of both the simulation study and
the case study using bird counts in mid-Scandinavia are presented. The paper finishes in
Sect. 6 with the discussion of the results and concluding remarks.

2. BIRD MONITORING SURVEYS DATA

2.1. TOV-E AND BBS DATA

The Norwegian common bird monitoring scheme (TOV-E), coordinated by the Nor-
wegian Institute for Nature Research (NINA) and the Norwegian Ornithological Society
(NOF) since 2006, was established to monitor population variation for common breeding
terrestrial birds on a national scale in a representative way. Surveys (i.e., count of pairs of
birds of all observed species) are carried out by experienced ornithologists that follow a
standardized protocol (Kålås 2002). Each census route (n = 492) contains between 12 and
20 (average = 18.8) point counts 300 m apart describing a square (see Fig. 1) with side =
1.5 km (deviation of this shape are allowed and recorded when the geographic/topographic
conditions do not allow the observer to walk, e.g., sea/lakes, glaciers, rough mountainous
terrain). A total of 229 species are heard or seen at the entirety of the point counts of TOV-E
during 5 minutes. Approximately 121 of the species are less abundant and/or difficult to
detect, so observers are asked to record these species during a line transect between point
counts (see Fig. 1—figure with the configuration of a census site with the twenty points).
A random selection of 370 census routes (out of a total of 492 routes across Norway) is
visited once a year during the period 20th May to 10th June. TOV-E is designed to cover all
relevant habitats throughout the altitudinal and latitudinal gradient in Norway and reports
'pairs of individuals' as sampling unit. The Swedish breeding bird survey (hereafter BBS)
has been coordinated by Lund University since 1996 and consists of 716 fixed sites across
Sweden within a 25-km grid (one route per grid cell, see Lindstrom et al. 2013). These sites
are surveyed once a year between mid-May and mid-June (the breeding period for most
bird species in Sweden) though not all sites are surveyed every year (mean = 353 sites per
year). The 25-km grid makes sure that the habitats of Sweden are monitored in proportion to
their abundance in the country as well as the entire altitudinal and latitudinal gradient where
birds are present. At each site, the observer walks an 8-km transect describing a $2 \times 2 \, \text{km}$
square and records all bird species heard and/or seen within 8 h. In addition, the observer
has eight 5-min point counts where all birds seen or heard must also be recorded. The point
counts take place at each of the corners of the square and at the middle point of the transect
(see Fig. 1). Of the circa 250 species breeding in Sweden, 244 are reported in BBS, thus
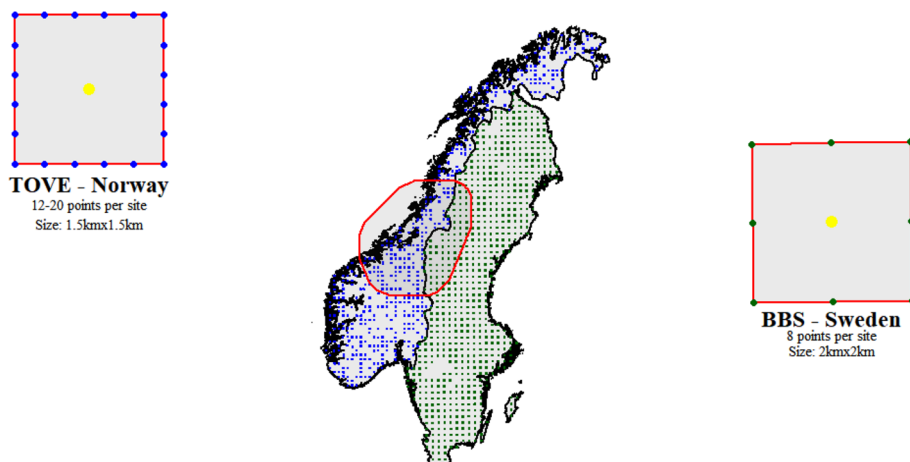
Figure 1. Spatial location of census sites and sampling points and line transects according to each sampling protocol. Left: graphical display of sampling protocol of TOV-E census. Blue points: 20 locations for point counts (the number of points vary between 12 and 20 in different sites). Red lines: line transects. Yellow point: centroid associated with each census site (see Sect. 2.2). Center: spatial distribution of census sites across Norway (blue sites) and Sweden (green sites). The red polygon represents the study area described in Sect. 2.1. Right: graphical display of sampling protocol of BBS census. Green points: 8 locations for point counts. Red lines: line transects. Yellow point: centroid associated with each census site (see Sect. 2.2).

ensuring a good coverage of the breeding birds (Lindstrom et al. 2013). The BBS reports 'individuals' as sampling unit, which differs from TOV-E's reporting unit (pairs; see above).

Although these monitoring programs are designed to cover a large part of both countries (Fig. 1), for our case study, we only selected census sites that lie within a polygon defined to produce an approximation of a Gaussian Random Field and make inference about a point pattern in Trøndelag Country, central Norway (see red polygon in Fig. 1, (Lindgren et al. 2011; Simpson et al. 2016)). This polygon covers a total area of 173.634 km^2 and contains 113 census sites in Norway and 70 in Sweden. The main motivation to reduce the study region from the entire country to a smaller area (defined by the polygon) was strictly computational and for an easier compilation of covariate information. In addition, this region, which is basically within Trøndelag County in central Norway, is largely representative of habitat types, topography and biodiversity found elsewhere in Norway.

2.2. Exploratory Analysis

Our main goal was to develop and validate a new modeling framework to integrate abundance data from standardized monitoring schemes with different sampling protocols. Such a framework can ultimately be used, for example, to detect hotspots of abundance of birds, as in the case we illustrate here (note: we are not interested in the distribution of particular species, but in the distribution of total abundance of birds regardless of the species). In other words, we apply our modeling framework to produce maps of total abundance of birds based on count data from multiple sources—information gathered as part of standardized national bird monitoring schemes in Norway and Sweden that differ in the sampling protocols. The

data preparation consisted in averaging across all years (2006–2019) the total count of all individuals (regardless of the species) found at each survey site. That is, we first added up the counts of all individual birds recorded in the points or lines of a given census site and assigned this total count of individuals (regardless of the species present) to the site's centroid (see Fig. 1) so that each census site will have one single value of total abundance of birds per year. Next, for each site, we averaged the yearly total abundance of birds across all years that the site was sampled (note: not all sites are censused every year) in the period between 2006 and 2019, so that we ended up with one single value of total abundance of birds per site (temporal average). Although estimating single-species abundance and distribution maps are commonly used to inform about species of conservation concern, here we wanted to report the total abundance of birds across the region (note: our methodology can also be used to estimate single-species abundances). Estimating total abundance of individuals across a region (as opposed to single-species abundance) has clear implications in spatial conservation planning and prioritization (Lehtomäki and Moilanen 2013). For example, De Lucas et al. (2008) estimated the total abundance of raptors in a region to assess the impacts of wind farms on this group of birds. Lindström et al. (2019) attempted to estimate total density of wading birds across Fennoscandia by combining count data from Norway, Sweden and Finland. However, they did not account for many differences in the sampling protocols. Our modeling framework thus can be applied to account for such differences. Another example of potential use of our method is to get more robust estimates of total abundance of birds to inform authorities and stakeholders where powerlines (Bevanger et al. 2014) or wind farms (De Lucas et al. 2008) may cause large mortality rates. Although here we present a simplified and more generic analysis (all species have weight = 1, and thus their abundance has the same influence in the resulting map), each species abundance can be multiplied (weighted) by a factor relative to their sensitivity to e.g., powerlines (D'Amico et al. 2019) so that the resulting map will highlight total abundance hotspots in relation to their sensitivity to the particular issue. Since we include data from both Norway and Swe-
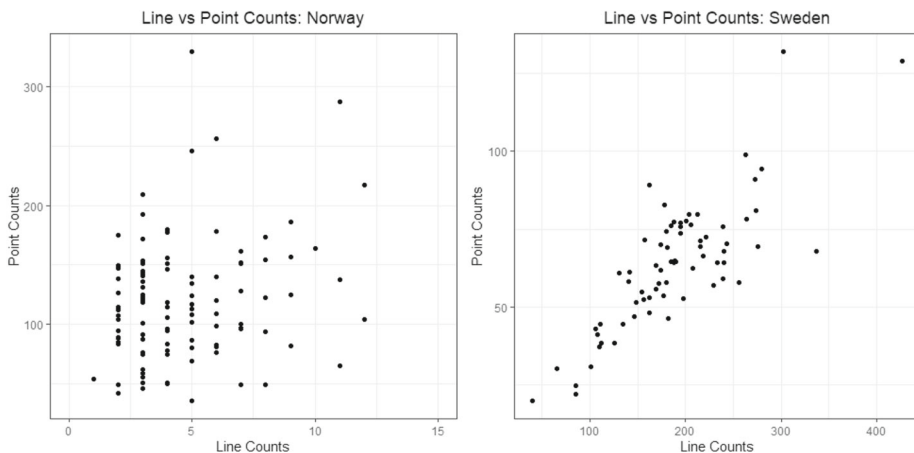

Figure 2.   Scatterplots of line vs point counts in Norway (number of pairs, left) and Sweden (number of individuals, right) .

den, we explore how the relation of point and line counts differ between surveys from both countries. In Fig. 2, we display a scatterplot with the points and line counts at each of the TOV-E (n=113) and BBS (n=70) sites.

These scatterplots show a linear relation between point and line counts in Sweden, whereas in Norway there is no clear linear association between the counts in points and lines. This is somehow expected due to the census design in Norway, where the line counts are meant to record a reduced subset of species compared to the point counts. This is a common issue highlighted by Buckland and Johnston (2017) and is often found in many countries when certain species are monitored with special censuses in addition to the general monitoring scheme. Therefore, this is not only an issue when integrating between-countries datasets (e.g., to increase the geographical extent), but also within-country datasets (to increase the representability and number of data points).

2.3. Explanatory variables

In our case study, we want to apply our new methodology not only to estimate total abundance of birds, but also to produce interpretable estimates of ecological factors associated with it across the region. We have selected three candidate ecological factors that are commonly used in SDMs to explain distribution of birds (e.g. Bradter et al. 2021; Lissovsky et al. 2021; Soultan et al. 2022): (i) climatic variables—temperature (average daily temperature from April to July over 2006–2019, downloaded from seNorge.no) and precipitation (average daily precipitation from April to July over 2006–2019, downloaded from seNorge.no), (ii) topography - elevation (Digital Elevation Model at a 10m resolution, DEM10, downloaded from https://kartkatalog.geonorge.no/), and (iii) the land cover surrounding each location expressed as the percentage of each of the following six land covers (urban, mountains, rocky area, water body, forest, and open area) in a square neighborhood of 2km × 2km. Land cover information was depicted from the N50 layer (downloaded from https://kartkatalog.geonorge.no/). All rasters files have resolution of 1km × 1km (the elevation data from DEM10 was aggregated to this resolution prior analysis) and are shown in the Supplementary Information. As a first stage of model selection, we computed the correlation coefficient between all the candidate covariates on a fine grid of about 600.000 points. Only one variable in those pairs with $|\rho| > 0.7$ was left as a candidate. Those pairs with high correlation were: 1) elevation and temperature ($\rho = -0.81$). Temperature was discarded; 2) % of open area and % of forest ($\rho = -0.83$). % of open area was discarded.

3. MODELING AND INFERENCE APPROACH

The specification of our models relies on the assumption that our four sources of observations are obtained from a common underlying ecological process (Miller et al. 2019). This assumption arguably makes sense if we consider the fact that national borders of neighboring countries are not, in general, a key factor for natural changes in biodiversity, although there might be slight differences in conservation policies and governance. Hence, we can assume that a common nonzero mean Gaussian Random Field (GRF) is involved in the

generation of the number of individuals at each census site. However, the two different
sampling protocols (points and lines), which also differ between the two countries (com-
plementary surveys in Norway and independent surveys in Sweden), result in four groups
of counts observed. Moreover, TOV-E counts (Norway) are reported as 'number of pairs'
of each species, whereas BBS counts (Sweden) are reported as 'number of individuals'
of each species. Therefore, direct inference and comparisons between these four response
variables should be made with caution. The true total bird counts random variable, $Y_{\text{true}}(\mathbf{s})$
with $\mathbf{s} \in D \subset \mathbb{R}^2$, is assumed to follow a Poisson distribution with expected value $\lambda_{\text{true}}(\mathbf{s})$,
expressed as

$$\log(\lambda_{\text{true}}(\mathbf{s})) = X^T(\mathbf{s})\beta + \omega_1(\mathbf{s}) \tag{1}$$

with $X^T(\mathbf{s})$ a set of spatial covariates and $\omega_1(\mathbf{s})$ a zero-mean GRF that aims at accounting
for residual spatial dependency. Both $X^T(\mathbf{s})$ and $\omega_1(\mathbf{s})$ can include well-established factors
that influence variation in the total abundance of birds; in our case study these factors are
precipitation and elevation. We assume a Matérn covariance function for $\omega_1(\mathbf{s})$

$$\frac{\sigma^2}{\Gamma(\nu)2^{\nu-1}}(\kappa\|s_i - s_j\|)^\nu K_\nu(\kappa\|s_i - s_j\|) \tag{2}$$

with $\|s_i - s_j\|$ the Euclidean distance between two locations $s_i, s_j \in D$. σ^2 stands for
the marginal variance, and K_ν represents the modified Bessel function of the second kind
and order $\nu > 0$. ν is the parameter that determines the degree of smoothness of the process,
while $\kappa > 0$ is a scaling parameter. For $\omega_1(\mathbf{s})$, let $\kappa = \kappa_1, \nu = \nu_1$ and $\sigma^2 = \sigma_1^2$. We assume
that the observed counts for each sampling protocol are realizations of four random variables
conditionally independent given the true abundance, $\lambda_{\text{true}}(\mathbf{s})$. That is, we assume the four
groups of observed counts are realizations of the Poisson random variables:

$$Y_1(\mathbf{s}) \sim \text{Poisson}(\lambda_1(\mathbf{s})) \quad \text{(Point counts in Norway)}$$
$$Y_2(\mathbf{s}) \sim \text{Poisson}(\lambda_2(\mathbf{s})) \quad \text{(Line counts in Norway)}$$
$$Y_3(\mathbf{s}) \sim \text{Poisson}(\lambda_3(\mathbf{s})) \quad \text{(Point counts in Sweden)}$$
$$Y_4(\mathbf{s}) \sim \text{Poisson}(\lambda_4(\mathbf{s})) \quad \text{(Line counts in Sweden)}$$

where $\lambda_j(\mathbf{s})$, $j = \{1, 2, 3, 4\}$ are the expected values of the random variables $Y_j(\mathbf{s})$.
Additionally, we assume $Y_1(\mathbf{s}) + Y_2(\mathbf{s}) \approx Y_{\text{NO}}(\mathbf{s})$ as a proxy for total abundance since the
line transects are complementary to the point counts in Norway. This assumption does not
hold for Sweden since, as mentioned in Sect. 1, line transects and point counts are regarded
as two different independent censuses. In case we wanted to suggest a proxy for the total
abundance in Sweden using $Y_3(\mathbf{s})$ and $Y_4(\mathbf{s})$, we would need to account for a potential overlap
(double counting) between the counts observed in points and line transects. Given that we
assume a common latent process underlying all the observed abundances, $Y_1(\mathbf{s}) + Y_2(\mathbf{s})$
works also as a proxy for total abundance of birds in Sweden. This variable is used to
produce the predicted total abundance of birds in Sect. 3. Our final assumption is that there

are no differences in observer skills between countries since the census are performed by experienced ornithologists.

3.1. Models

In this section, we introduce three model specifications for integrating data from the four sampling protocols introduced in Sect. 2. Model 1 (see Sect. 3.1.1) is a model that assumes a linear relation between the expected counts of the four sampling protocols. This is achieved by the introduction of a unique intercept for each sampling scheme. In Sect. 3.1.2, model 2 is presented. This model allows for a relaxation of the assumption of linear relation between expected counts by incorporating terms that allow to explain any deviation from this assumption through the GRF $\omega_1(\mathbf{s})$. Finally, model 3 (see Sect. 3.1.3) is introduced. This model adds a second GRF, $\omega_2(\mathbf{s})$, which aims to account for spatial sources of variation not accounted for in the other parts of the model and not explained by known covariates, (Simmonds et al. 2020; Selle et al. 2020). It is worth noting that as each of the models proposed depend on $\lambda_{\text{true}}(\mathbf{s})$, they explicitly account for the factors that influence the variation in abundance.

3.1.1. Model 1

Based on our exploratory analysis and the four sampling processes present in our dataset, in model 1 we assumed a linear relation between the expected values of the four random variables representing each sampling protocol and $\lambda_{\text{true}}(\mathbf{s})$. That is,

$$
\begin{aligned}
\lambda_1(\mathbf{s}) &= \zeta_1^* \cdot \lambda_{\text{true}}(\mathbf{s}); & \log(\zeta_1^*) &\sim N(0, \tau_1^*) \\
\lambda_2(\mathbf{s}) &= \zeta_2^* \cdot \lambda_{\text{true}}(\mathbf{s}); & \log(\zeta_2^*) &\sim N(0, \tau_2^*) \\
\lambda_3(\mathbf{s}) &= \zeta_3^* \cdot \lambda_{\text{true}}(\mathbf{s}); & \log(\zeta_3^*) &\sim N(0, \tau_3^*) \\
\lambda_4(\mathbf{s}) &= \zeta_4^* \cdot \lambda_{\text{true}}(\mathbf{s}); & \log(\zeta_4^*) &\sim N(0, \tau_4^*)
\end{aligned}
\tag{3}
$$

with $\zeta_j^* \geq 0$, $j = 1, \ldots, 4$ the factors that determine the association between the observed and the true counts for each protocol. In real-life problems, ζ_j^* can explain multiple sources of variation that are common to sampling of bird species such as observer differences, observed units, differences in detection probability, among others. The inclusion of this term is also useful to deal with overdispersion (Gomez-Rubio 2020), a common issue when working with count data. In order to avoid identifiability issues, we restate the model in (3) in terms of $\lambda_1(\mathbf{s})$. That is,

$$
\begin{aligned}
\lambda_2(\mathbf{s}) &= \zeta_2 \cdot \lambda_1(\mathbf{s}); & \log(\zeta_2) &\sim N(0, \tau_2) \\
\lambda_3(\mathbf{s}) &= \zeta_3 \cdot \lambda_1(\mathbf{s}); & \log(\zeta_3) &\sim N(0, \tau_3) \\
\lambda_4(\mathbf{s}) &= \zeta_4 \cdot \lambda_1(\mathbf{s}); & \log(\zeta_4) &\sim N(0, \tau_4)
\end{aligned}
\tag{4}
$$

where $\zeta_j \geq 0$ and $\zeta_j = \frac{\zeta_j^*}{\zeta_1^*}$, $j = \{2, 3, 4\}$.

3.1.2. Model 2

In model 2, we relax the assumption of linear relation between the expected value of the number of observed individuals with protocol j, $\lambda_j(\mathbf{s})$, and the true intensity, $\lambda_{\text{true}}(\mathbf{s})$, by including spatial varying terms $(\psi_j^* - 1) \cdot \omega_1(\mathbf{s})$, $j = \{1, 2, 3, 4\}$. These terms aim to explain any deviation from a linear relation between expected values as a function of a GRF $\omega_1(\mathbf{s})$. It is worth noting that model 1 (see above) is a special case of model 2 with $\psi_j^* = 1$. We define model 2 as:

$$\lambda_1(\mathbf{s}) = \zeta_1^* \cdot \lambda_{\text{true}}(\mathbf{s}) \cdot \exp\{(\psi_1^* - 1) \cdot \omega_1(\mathbf{s})\}; \quad \log(\zeta_1^*) \sim N(0, \tau_1^*)$$
$$\lambda_2(\mathbf{s}) = \zeta_2^* \cdot \lambda_{\text{true}}(\mathbf{s}) \cdot \exp\{(\psi_2^* - 1) \cdot \omega_1(\mathbf{s})\}; \quad \log(\zeta_2^*) \sim N(0, \tau_2^*)$$
$$\lambda_3(\mathbf{s}) = \zeta_3^* \cdot \lambda_{\text{true}}(\mathbf{s}) \cdot \exp\{(\psi_3^* - 1) \cdot \omega_1(\mathbf{s})\}; \quad \log(\zeta_3^*) \sim N(0, \tau_3^*)$$
$$\lambda_4(\mathbf{s}) = \zeta_4^* \cdot \lambda_{\text{true}}(\mathbf{s}) \cdot \exp\{(\psi_4^* - 1) \cdot \omega_1(\mathbf{s})\}; \quad \log(\zeta_4^*) \sim N(0, \tau_4^*) \tag{5}$$

Again, to avoid identifiability issues, we restate the model in (5) in terms of $\lambda_1(\mathbf{s})$ as:

$$\lambda_2(\mathbf{s}) = \zeta_2 \cdot \lambda_1(\mathbf{s}) \cdot \exp\{(\psi_2 - 1) \cdot \omega_1(\mathbf{s})\}; \quad \log(\zeta_2) \sim N(0, \tau_2)$$
$$\lambda_3(\mathbf{s}) = \zeta_3 \cdot \lambda_1(\mathbf{s}) \cdot \exp\{(\psi_3 - 1) \cdot \omega_1(\mathbf{s})\}; \quad \log(\zeta_3) \sim N(0, \tau_3)$$
$$\lambda_4(\mathbf{s}) = \zeta_4 \cdot \lambda_1(\mathbf{s}) \cdot \exp\{(\psi_4 - 1) \cdot \omega_1(\mathbf{s})\}; \quad \log(\zeta_4) \sim N(0, \tau_4) \tag{6}$$

In the scales of the linear predictors in (5), $\psi_j = \psi_j^* - \psi_1^* + 1$, $j = \{2, 3, 4\}$ are scaling coefficients for the common GRF, $\omega_1(\mathbf{s})$, in each likelihood. They quantify to what extent the departure of the assumption of linearity is explained by $(\psi_j^* - 1) \cdot \omega_1(\mathbf{s})$. In real-life scenarios, this departure can be related with sources of variation with spatial structure such as differences in detectability, among others. Therefore, we would expect posterior densities for ψ_3 and ψ_4 to be around 1 in our case study, while for ψ_2 we expect different results because line and point counts in Norway do not seem to follow a linear relation (see Sect. 2; Fig. 2). Due to different characteristics of line transect surveys in Norway, we propose model 3.

3.1.3. Model 3

In addition to causing departure from a linear relation between true and observed counts, species detectability may also change with the census technique used (i.e., one of our data sources, the line transects in TOV-E, targeted only a subset of species as it is regarded as a complementary survey to the point counts). Hence, in model 3 we included a second GRF, $\omega_2(\mathbf{s})$ to try to account for the characteristics of this observation process. In case that no explanatory variable that explains the particular characteristics of the sampling protocol is available, a second GRF can be added as a way to account for them, (Simmonds et al. 2020). This is included as an additive term in the linear predictor, as follows:

$$\lambda_2(\mathbf{s}) = \zeta_2 \cdot \lambda_1(\mathbf{s}) \cdot \exp\{(\psi_2 - 1)\omega_1(\mathbf{s})\} \cdot \exp\{\omega_2(\mathbf{s})\}$$
$$\lambda_3(\mathbf{s}) = \zeta_3 \cdot \lambda_1(\mathbf{s}) \cdot \exp\{(\psi_3 - 1)\omega_1(\mathbf{s})\}$$
$$\lambda_4(\mathbf{s}) = \zeta_4 \cdot \lambda_1(\mathbf{s}) \cdot \exp\{(\psi_4 - 1)\omega_1(\mathbf{s})\} \tag{7}$$

We assume a Matérn covariance function as in (2) for $\omega_2(\mathbf{s})$, with parameters $\kappa = \kappa_2$, $\nu = \nu_2$ and $\sigma^2 = \sigma_2^2$

3.1.4. Prior Specification

For the GRFs $\omega_k(\mathbf{s})$, $k = \{1, 2\}$, the parameters ν_k in the Matérn covariance function are fixed to be 1. The interest is put on the spatial ranges ρ_k, and on the standard deviation of the GRFs, σ_k. ρ_k are related to κ_k through $\rho_k = \sqrt{8}/\kappa_k$. The prior distributions of these two parameters are specified by making use of Penalized Complexity (PC) priors, (Fuglstad et al. 2019). In this case, we set $P(\rho_1 < 20000) = 0.1$ and $P(\sigma_1 > 1) = 0.1$ for $\omega_1(\mathbf{s})$, while $P(\rho_2 < 2000) = 0.1$ and $P(\sigma_2 > 3) = 0.1$ for $\omega_2(\mathbf{s})$. This means, for example, that under this prior specification, a standard deviation greater than 1 is regarded as large, while a spatial range below 20 kilometers is considered unlikely for $\omega_1(\mathbf{s})$. The parameters in $\boldsymbol{\beta}$ have Normal prior with mean 0 and precision 0.01. Let $\log(\zeta_j) \sim N(0, \tau_j)$, $j = \{2, 3, 4\}$, where the logarithm of each τ_j has a log-Gamma prior with parameters 1 and 0.00005. For the parameters ψ_j, $j = 2, 3, 4$ in models 2 and 3, we set a normal prior with mean 1 and precision 0.1. We have now defined a group of three candidate models. In the upcoming subsections, we introduce the methodological approach for fitting them and for selecting a model that suits best for our problem.

3.2. Inference and Computational Approach

The models introduced in Sect. 3.1 were fitted making use of the Integrated Nested Laplace Approximation (INLA), (Rue et al. 2009) and the Stochastic Partial Differential Equation (SPDE) approach (Lindgren et al. 2011). INLA is a faster alternative to Monte Carlo Markov Chains (MCMC) for performing Bayesian inference for latent Gaussian models. INLA aims at producing a numerical approximation of the marginal posterior distribution of the parameters and hyperparameters of the model. Further details can be found in Rue et al. (2009) and Blangiardo and Cameletti (2015). Since we deal with continuous spatial processes in our models, the SPDE approach emerges as an efficient representation of $\omega_1(\mathbf{s})$ and $\omega_2(\mathbf{s})$. It is based on the solution of a SPDE which can be approximated through a basis function representation defined on a triangulation of the spatial domain. More details are available in Lindgren et al. (2011) and Blangiardo and Cameletti (2015).

3.3. Assumptions and Possible Extensions

This new modeling framework is developed to integrate count data collected in designed surveys that follow different standardized protocols. Particularly, in the case study presented here, the bird surveys introduced in Sect. 2 are designed to minimize biases due to variation in the time of sampling or observer expertise. For this reason, the models presented in our case study assume, in principle, that these external sources of variation that could affect the observation process are constant across sites or negligible. However, these models are flexible enough to explicitly account for factors that may affect the observation process of each sampling protocol, and can thus be accounted for. There may be, however, other potential sources of variation when working with monitoring data, which also depend on the

taxon being surveyed. Hence, as mentioned in Sect. 3, our method includes relevant terms for quantifying the effect of potential sources of noise in the observation process. Our models incorporate the terms ζ to explain what proportion of the true abundance is explained by each of the observation processes. That is, ζ_j quantifies the effect of each sampling protocol on the observed abundances. This effect comprises sources of variation such as differences in the observed units, differences in detectability, and potential differences in the expertise of the observers. In many real-life scenarios, these terms do not provide enough quantification of the effect of the sampling protocols as there are sources of variation in the sampling process that have spatial variation that cannot be summarized in one term. Therefore, the Gaussian Random Field that drives the true abundance (in our case study, the total abundance of birds) or a second GRF is also used to account for sources of variation that have a spatial behavior. This modeling framework also allows to explicitly account for factors that affect the observation process of each sampling protocol. To show how this can be done, we take model 2 as our reference to explicitly account for a factor that influences the observed number of individuals. We now assume that unlike our case study, there are several factors affecting the observed total abundance of birds. As seen in equation (6) in Sect. 3.1.2, the term $\zeta_j \cdot \exp\left\{(\psi_j - 1)\omega_1(\mathbf{s})\right\}$ accounts for the effect the sampling protocol j has on the observed abundance. In addition to the spatial effect driven by $\omega_1(\mathbf{s})$, the term ζ_j can be further explained, for example, by a fixed effect z as follows:

$$\zeta_j = \alpha_{0j} + \alpha_{1j}z \tag{8}$$

This is a straightforward way to explicitly account for multiple factors that may influence the observation process of the sampling protocol j. Factors with a spatial or temporal structure can be accounted for through random effects with these structures. Given the additional parameters to be estimated and the increased complexity of the model when the effect of these factors is accounted for explicitly, structural identifiability issues may arise. Therefore, in order to overcome these issues, it is recommended to constrain the parameters in (8). This can be achieved by either having additional data that inform on these factors or informative prior information of the parameters involved in (8). Acquiring additional data to account for factors that affect the observation process of each sampling protocol might be possible by integrating data, for example, from schemes with sampling protocols designed to gather information on species detection probabilities through repeated visits to the sites or distance sampling (Järvinen and Väisänen 1983; Miller et al. 2019). In our case study, the temporal variation in birds is not considered to compute the total abundance of birds across the study region. Rather, this temporal variation is removed by averaging the total count of birds at each site over the 14 years (2006–2019). This is also a convenient assumption as we do not have information (counts) at every census site every year (i.e., not all sites are surveyed every year). Furthermore, we believe that the overall state of important sites for birds has remained similar in the past 14 years (i.e. bird-rich areas in 2006, at the beginning of the monitoring scheme are still bird-rich areas in 2019, even if the species composition might have changed slightly).

3.4. Model Assessment

In order to assess and compare competing models such as the ones we are fitting in upcoming sections, we employed the Deviance Information Criterion (DIC), (Spiegelhalter et al. 2002), the Watanabe–Akaike Information Criterion (WAIC), (Watanabe 2010), the logarithm of the pseudo marginal likelihood (LPML) (Blangiardo and Cameletti 2015) and the Continuous Rank Probability Score (CRPS) (Gneiting and Raftery 2007).

DIC makes use of the deviance of the model

$$D(\theta) = -2\log(p(\mathbf{y}|\boldsymbol{\theta}))$$

to compute the posterior mean deviance $\bar{D} = E_{\boldsymbol{\theta}|\mathbf{y}}(D(\boldsymbol{\theta}))$. In order to penalize the complexity of the model, the effective number of parameters

$$p_D = E_{\boldsymbol{\theta}|\mathbf{y}}(D(\boldsymbol{\theta})) - D(E_{\boldsymbol{\theta}|\mathbf{y}}(\boldsymbol{\theta})) = \bar{D} - D(\bar{\boldsymbol{\theta}})$$

is added to $\bar{D}$. Thus,

$$\text{DIC} = \bar{D} + p_D.$$

The Watanabe–Akaike Information Criterion is based on the posterior predictive density, which makes it preferable to the Akaike and the deviance information criteria, since according to Gelman et al. (2014) it averages over the posterior distribution rather than conditioning on a point estimate. It is empirically computed as

$$-2\left[\sum_{i=1}^{n}\log\left(\frac{1}{S}\sum_{s=1}^{S}p(y_i|\theta^s)\right) + \sum_{i=1}^{n}V_{s=1}^{S}(\log p(y_i|\theta^s))\right]$$

with θ^s a sample of the posterior distribution and $V_{s=1}^{S}$ the sample variance. Another criterion to compare the models is LMPL, defined as:

$$\text{LPML} = \sum_{i=1}^{n}\log(\text{CPO}_i)$$

It depends on CPO_i, the Conditional Predictive Ordinate at location $\mathbf{s}_i$, (Pettit 1990), a measure that assesses the model performance by means of leave-one-out cross validation. It is defined as:

$$\text{CPO}_i = p(y_i^*|y_f)$$

with y_i^* the prediction of y at location $\mathbf{s}_i$ and $y_f = y_{-i}$. Lastly, we will compare the predictive performance of our models using the Continuous Rank Probability Score (CRPS). It makes possible to compare the estimated posterior mean and our observed values while accounting for the uncertainty of the estimation, (Gneiting and Raftery 2007; Selle et al. 2019). It is defined as:

$$\mathrm{CRPS}(F, y) = \int_{-\infty}^{\infty} (F(u) - 1\{y \le u\})^2 du$$

with F, the cumulative distribution of the estimated posterior mean, and y is the observed value. The smaller CRPS is, the closer the estimated value is to the observed one.

4. SIMULATION STUDIES

We set up three simulation studies based on the case study of total abundance of birds in mid-Scandinavia that allow us to assess the performance of the models proposed in Sect. 3, when the true data generating model either assume linear relation between the counts (Scenario 1), deviate from this assumption due to some spatial factor explained by a GRF (Scenario 2) or when one group of observed counts is considerably affected by additional spatial sources of variation (Scenario 3). We used the same sites as the observations in the TOV-E and BBS surveys (Fig. 1). To start, we simulated the true intensity, $\lambda_{true}(\mathbf{s})$ as:

$$\log(\lambda_{\mathrm{true}}(\mathbf{s})) = \beta_0 + \beta_1 \mathrm{PREC}(\mathbf{s}) + \omega_1(\mathbf{s})$$

with $\mathrm{PREC}(\mathbf{s})$, the precipitation at location $\mathbf{s}$ in the study region (see Figure S.1.), and $\omega_1(\mathbf{s})$ a GRF with range $\rho = 15 \mathrm{km}$ and $\sigma^2 = 0.14$. Further, we specified $\beta_0 = 4.70$ and $\beta_1 = -0.20$. These values were chosen based on the posterior marginal distribution of these parameters in the real-data application. Next, we simulated observations representing the surveys, i.e., using four different Poisson models with parameters $\lambda_j(\mathbf{s})$, $j = \{1, \ldots, 4\}$. Table 1 summarizes the two simulation scenarios proposed for $\lambda_j(\mathbf{s})$

For each scenario, we simulated 100 datasets with $\zeta_1^* = 0.91$, $\zeta_2^* = 0.04$, $\zeta_3^* = 0.57$ and $\zeta_4^* = 1.72$. While we assume a linear relation between $\lambda_j(\mathbf{s})$ and $\lambda_{true}(\mathbf{s})$ in Scenario 1, in Scenarios 2 and 3 the relation between $\lambda_j(\mathbf{s})$ and $\lambda_{true}(\mathbf{s})$ is assumed to follow (5) with $\psi_1^* = 1$, $\psi_2^* = 1.57$, $\psi_3^* = 1.09$ and $\psi_4^* = 1.21$. These settings are based on the posterior marginal distribution of the parameters in the real data case study (presented in Sect. 5.2). The three simulation scenarios closely mimicked real data application by making two of the simulated counts only observed in Norway and the other two only observed in Sweden. For each simulated dataset, we fitted the three models proposed in Sect. 3. A second group of simulation scenarios was proposed by taking more extreme values of the posterior marginal distributions. The results and more details on this simulation scenario are discussed in Sect. 5.1 and the supplementary information.

Table 1. Simulation scenarios

Scenario	Simulated $\lambda_j(\mathbf{s})$
1	$\lambda_j(s) = \zeta_j^* \cdot \lambda_{true}(\mathbf{s})$
2	$\lambda_j(s) = \zeta_j^* \cdot \lambda_{true}(\mathbf{s}) \cdot \exp((\psi_j^* - 1) \cdot \omega_1(\mathbf{s}))$
3	$\lambda_j(s) = \zeta_j^* \cdot \lambda_{true}(\mathbf{s}) \cdot \exp((\psi_j^* - 1) \cdot \omega_1(\mathbf{s}))$; $j = \{1, 3, 4\}$
	$\lambda_2(s) = \zeta_2^* \cdot \lambda_{true}(\mathbf{s}) \cdot \exp((\psi_2^* - 1) \cdot \omega_1(\mathbf{s}) + \omega_2(\mathbf{s}))$

To assess the performance of each model in each scenario, we simulated 10000 realizations $\{\theta^p_{jkl}\}$, $j =, 1 \ldots, 10000$, from the posterior distribution of each parameter θ for dataset $k = 1, \ldots, 100$ in scenario $l = 1, 2, 3$. Thus, the mean bias and the Root Mean Square Error (RMSE) for dataset k in scenario l are computed as:

$$\text{bias}_{kl} = \frac{1}{10000} \sum_{j=1}^{10000} \left(\theta^p_{jkl} - \tilde{\theta} \right)$$

$$\text{RMSE}_{kl} = \sqrt{\frac{1}{10000} \sum_{j=1}^{10000} \left(\theta^p_{jkl} - \tilde{\theta} \right)^2}$$

with $\tilde{\theta}$ the true value of the parameter θ.

5. RESULTS

5.1. SIMULATION STUDIES

The 100 datasets generated in each of the proposed scenarios were fitted using the three proposed models in Sect. 3 and the results summarized here using the measures of performance introduced in Sect. 4. We only show the mean bias and RMSE for the parameters ζ_2^*, ζ_3^* and ζ_4^* as they are key to understand how different response variables interact with each other (Fig. 3).
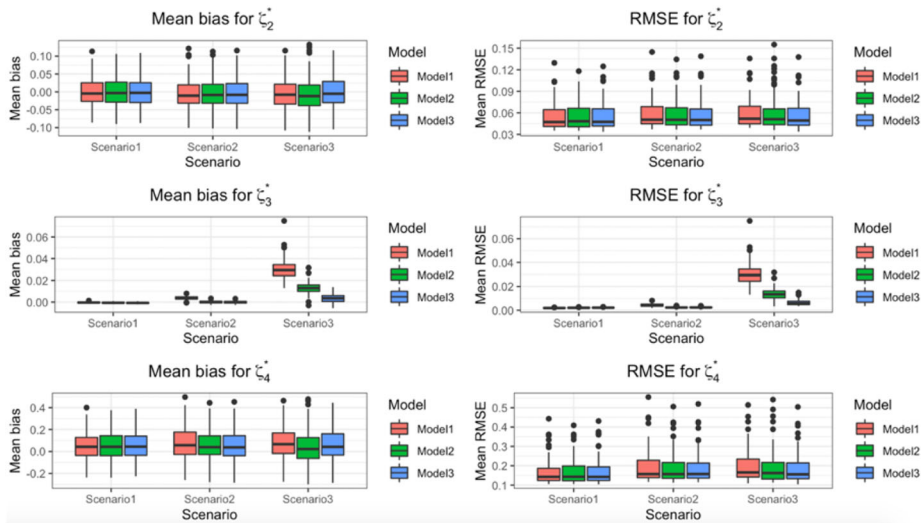

Figure 3.   Mean bias (left) and RMSE (right) for parameters ζ_2^* (upper panels), ζ_3^* (central panels) and ζ_4^* (lower panels) for each model in simulation scenario 1 (assumption of linear relationship between expected abundances), scenario 2 (non-linear relation between expected abundances explained by $\omega_1(\mathbf{s})$) and scenario 3 (an extra spatial source of variation affecting only one of the groups of observed counts).

Table 2. Mean bias and RMSE for parameters β_0, β_1, ρ and σ in simulation scenario 1 (assumption of linear relationship between expected abundances), scenario 2 (non-linear relation between expected abundances explained by $\omega_1(\mathbf{s})$) and scenario 3 (an extra spatial source of variation affecting only one of the groups of observed counts)

Scenario	Model	β_0		β_1		ρ (km)		σ	
		Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
1	1	-0.112	0.120	$1.62 \cdot 10^{-3}$	0.040	-1.567	4.729	0.076	0.097
		(0.037)	(0.035)	(0.028)	(0.011)	(3.982)	(1.946)	(0.080)	(0.076)
	2	-0.116	0.125	$2.74 \cdot 10^{-3}$	0.043	-1.380	4.812	0.115	0.129
		(0.036)	(0.034)	(0.028)	(0.011)	(4.330)	(2.267)	(0.096)	(0.099)
	3	-0.120	0.129	$1.48 \cdot 10^{-4}$	0.044	-1.497	4.762	0.122	0.135
		(0.038)	(0.035)	(0.030)	(0.011)	(3.978)	(2.021)	(0.088)	(0.093)
2	1	-0.112	0.119	$1.37 \cdot 10^{-3}$	0.040	-1.132	4.681	0.066	0.089
		(0.037)	(0.035)	(0.028)	(0.011)	(4.152)	(2.057)	(0.070)	(0.065)
	2	-0.111	0.120	$1.05 \cdot 10^{-4}$	0.038	-0.880	4.704	0.069	0.091
		(0.036)	(0.034)	(0.024)	(0.009)	(4.160)	(2.148)	(0.070)	(0.065)
	3	-0.104	0.113	$-1.92 \cdot 10^{-3}$	0.048	-0.952	4.629	0.058	0.082
		(0.038)	(0.035)	(0.049)	(0.025)	(4.013)	(1.981)	(0.067)	(0.060)
3	1	-0.112	0.120	$1.42 \cdot 10^{-3}$	0.040	-1.056	4.652	0.063	0.087
		(0.037)	(0.035)	(0.028)	(0.011)	(3.927)	(1.972)	(0.069)	(0.065)
	2	-0.111	0.120	$5.06 \cdot 10^{-5}$	0.038	-0.596	4.783	0.069	0.089
		(0.036)	(0.034)	(0.024)	(0.009)	(4.288)	(2.139)	(0.070)	(0.065)
	3	-0.112	0.119	$-2.05 \cdot 10^{-4}$	0.040	-1.837	5.024	0.088	0.107
		(0.037)	(0.035)	(0.028)	(0.011)	(4.240)	(2.064)	(0.079)	(0.076)

In parentheses, the standard error of each performance measurement

Figure 3 shows that the estimation of the proportional relation between the four likelihoods performed similarly for the three models when the truth is that the four likelihoods are linearly related (Scenario 1). Model 1 (which assumes linear relationship between the expected counts) performed, as expected, slightly better than the other two models as this is the model that generated the datasets. However, when we introduced some deviation from the assumption of linearity in our data generating process (Scenario 2), model 1 underperformed relative to the other two models. This is true for the three parameters of interest (Fig. 3). Models 2 and 3 performed better in terms of bias and RMSE, whereas the estimates produced by model 1 were biased and showed higher variability. Lastly, when an additional source of variation affected only one of the likelihoods (Scenario 3), the three models performed similarly as in Scenario 2, except for the hyperparameter ζ_3^*, which is part of the likelihood affected by the extra source of variation. For this hyperparameter, the differences in performance between the three models increased considerably as model 3 produced less biased and variable estimates of this hyperparameter.

Our results show that there are only marginal differences in the fixed effects β_0 and β_1 between the three models in all the scenarios. However, larger differences are observed for the hyperparameters of $\omega_1(\mathbf{s})$. For example, in the three scenarios the bias of ρ was smaller for model 2 compared to the other two models, but at the same time it produced estimates of ρ with larger RMSE than the other two models. In this simulation study, we have also explored the selection of the best model according to the comparison criteria DIC, WAIC

Figure 4. Differences in DIC, WAIC and LMPL between the model that generated the observed counts in each simulation scenario (Scenario 1, generated according to model 1; scenario 2, generated according to model 2 and scenario 3, generated according to model 3) and the other two models proposed in Sect. 3.

and LMPL (See Sect. 3.4). For each scenario, we computed the differences in each criterion between the model that generated the 100 datasets of the scenario and the other two models. The summaries of these differences are displayed in Fig. 4.

Figure 4 shows small differences in DIC and WAIC between the three models when model 1 generates the observed counts (Scenario 1). In Scenario 1, the predictive performance, measured by LMPL, was similar for model 1 (the one that generated that data) and model 2, while model 3 underperformed. In Scenario 2, model 2 (generating model) and model 3 performed similarly based on all performance comparisons, but model 1 underperformed considerably. In scenario 3, where the observed counts are generated according to a more complex specification (i.e., one sampling protocol is affected by an additional source of variation), model 3 had better goodness of fit and predictive performance with large differences in DIC, WAIC and LMPL with respect to the other two models. The difference in performance between models increases as the complexity of the data generating process increases (Fig. 4).

The results for the second group of simulations can be found in the Supplementary Information. Results match those obtained with the first group of simulations above. In Scenario 1, all three models perform similarly. As in the first simulation study, when the complexity of the model that generates the data increases, models 2 and 3 outperform model 1. Nevertheless, unlike in the first simulation study, model 2 outperformed model 3 in Scenario 3 (generated by model 3) for ζ_3 as it produced less biased estimates.

5.2. Results of the Case Study on Total Abundance of Birds in Mid-Scandinavia

We fitted our three models (see Sect. 3) to count data from the common bird monitoring schemes in Norway and Sweden (see Sect. 2) to estimate total abundance of birds across mid-Scandinavia with precipitation and elevation as explanatory variables. These two were selected from all the variables considered *a priori*, as it was the subset of candidate variables that produced the best results in terms of goodness of fit (see Supplementary Information for an overview of the performance of other competing models). The most demanding model in terms of computation time was model 3, which run in 60 seconds. In Table 3, we report the posterior mean, standard deviation and quartiles of the most relevant parameters from the three models.

Table 3 shows the associations between precipitation (PREC) and elevation (ELEV) with the expected counts are negative for all the models. The posterior means of the parameters of these two variables have small differences, model 2 estimated stronger association of the explanatory variables (precipitation and elevation) and the response variable (total abundance of birds). The posterior summaries of PREC and ELEV suggest that those locations with higher levels of precipitation and high elevation are expected to have lower total bird counts. The variability and range of the Gaussian field have right skewed posterior distributions based on their posterior medians and means.

Figure 5 and Table 3 show that the posterior densities of ζ_2 are different between models, with higher posterior mean for model 1 compared to the other models. This result agrees with the exploratory analysis of Sect. 2, which suggested the necessity of specifying a relaxed linear relationship between the line and point counts in Norway (linearity was met in Sweden, but not in Norway, see Fig. 2). However, the posterior densities of ζ_3 and ζ_4 are almost identical for models 1 and 3, whereas model 2 estimated posterior distributions for ζ_3 and ζ_4 that are shifted toward lower values (Fig. 5). Large differences in the posterior mean of ψ_2 in models 2 and 3 are observed when $\omega_2(\mathbf{s})$ is introduced to account for the particularities of the sampling protocol of the line counts in Norway (i.e., in general terms, to account for added complexity due to one of the data collecting protocols considered). While model 2 gives high prevalence to $\omega_1(\mathbf{s})$ (posterior mean of $\psi_2 = 1.90$) as determinant of the departure from linear association, model 3 reduces this prevalence (posterior mean of $\psi_2 = 0.63$). It arguably means that $\omega_2(\mathbf{s})$ accounts for what is particular of this sampling protocol (the added complexity) and what at the same time reduces the leverage of what is shared between this sampling protocol (the line transect in Norway in this case study) and the other protocols. We expect these differences in contribution of $\omega_1(\mathbf{s})$ across models to impact their predictive performance. In Figure S.2, we show the posterior mean of $Y_1(\mathbf{s}) + Y_2(\mathbf{s})$, understood as a proxy for the total abundance of birds in our study region (see Sect. 3). Given the high similarity across mid-Scandinavia, hereafter, we explore the differences in the predicted mean of $Y_1(\mathbf{s}) + Y_2(\mathbf{s})$ between the three models in a smaller sub-region (highlighted with a red square in Fig. 6), which encompasses the locations surrounding Trondheimsfjorden and the Norwegian Sea.

Our three models predicted high total bird counts along the eastern coast of Trondheimsfjorden and on the islands of Hitra and Frøya (Fig. S.9) and low counts at higher elevations

Table 3. Posterior mean, standard deviation and quartiles of the most relevant parameters of the models proposed in Sect. 3

			Model Model 1		
Parameter	Mean	SD	0.025q	0.50q	0.975q
Intercept	4.69	0.04	4.61	4.69	4.77
PREC	-0.12	0.04	-0.19	-0.12	-0.04
ELEV	-0.29	0.04	-0.38	-0.29	-0.21
ζ_2	0.05	0.00	0.04	0.05	0.05
ζ_3	0.51	0.03	0.45	0.51	0.57
ζ_4	1.50	0.09	1.33	1.50	1.68
ψ_2					
ψ_3					
ψ_4					
$\rho(m)$	$1.80 \cdot 10^4$	$4.00 \cdot 10^3$	$1.11 \cdot 10^4$	$1.77 \cdot 10^4$	$2.68 \cdot 10^4$
σ	0.36	0.02	0.32	0.36	0.41

			Model Model 2		
Parameter	Mean	SD	0.025q	0.50q	0.975q
Intercept	4.68	0.03	4.62	4.68	4.75
PREC	-0.20	0.03	-0.26	-0.20	-0.14
ELEV	-0.39	0.04	-0.46	-0.39	-0.32
ζ_2	0.04	0.00	0.04	0.04	0.05
ζ_3	0.48	0.03	0.43	0.48	0.54
ζ_4	1.42	0.08	1.27	1.42	1.58
ψ_2	1.86	0.14	1.59	1.86	2.13
ψ_3	1.26	0.13	1.00	1.26	1.52
ψ_4	1.30	0.12	1.07	1.30	1.54
$\rho(m)$	$1.80 \cdot 10^4$	$3.88 \cdot 10^3$	$1.17 \cdot 10^4$	$1.75 \cdot 10^4$	$2.69 \cdot 10^4$
σ	0.31	0.02	0.27	0.31	0.36

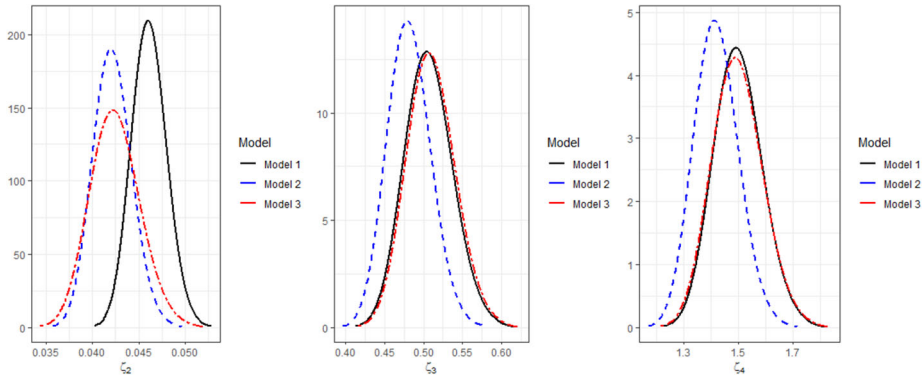|			Model Model 3		
Parameter	Mean	SD	0.025q	0.50q	0.975q
Intercept	4.69	0.04	4.61	4.69	4.77
PREC	-0.11	0.04	-0.18	-0.11	-0.04
ELEV	-0.27	0.04	-0.35	-0.27	-0.19
ζ_2	0.04	0.00	0.04	0.04	0.05
ζ_3	0.51	0.03	0.45	0.51	0.57
ζ_4	1.50	0.09	1.32	1.49	1.69
ψ_2	0.61	0.16	0.30	0.61	0.91
ψ_3	1.09	0.12	0.86	1.09	1.34
ψ_4	1.18	0.12	0.96	1.18	1.42
$\rho(m)$	$2.01 \cdot 10^4$	$4.12 \cdot 10^3$	$1.29 \cdot 10^4$	$1.98 \cdot 10^4$	$2.90 \cdot 10^4$
σ	0.34	0.03	0.29	0.34	0.39

Figure 5.   Posterior densities of ζ_2 (left), ζ_3 (center) and ζ_4 (right) for each model.
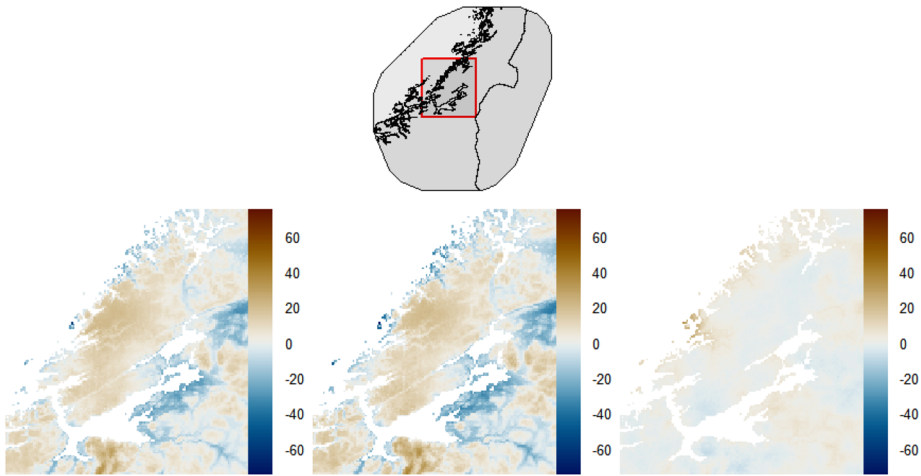

Figure 6.   Top(small): Study region with the red square that encloses the zone chosen for analyzing differences between models. Bottom: differences in the posterior mean of $Y_1(\mathbf{s}) + Y_2(\mathbf{s})$ (i.e. total abundance of birds) between: model 1 - model 2 (left), model 3 - model 2 (center) and model 1 - model 3 (right).

such as in the mountainous in the southwest and the north of the study region (Fig S.9.). Model 2 estimates higher counts compared to the other two models along the fjord's coast (dark blue) and lower abundance inland (mainly in the mountains; light brown; Fig. 6). The differences in predicted counts between model 1 and model 3 are smaller (Fig. 6, right panel) compared to those with model 2. However, larger predicted counts are produced by model 3 around the island of Linesøya. Our modeling framework allows for computing the uncertainty of our predictions. Here, we assess this by computing the standard error of $Y_1(\mathbf{s}) + Y_2(\mathbf{s})$ (see Fig. 7 for the standard error of the sub-region highlighted in Fig. 6, and see Fig. S.10 for the standard errors across the entire study region).

The standard error of model 1 is larger than the other two models in most regions (see brown colors, left and right panels in Fig. 7). In the zones with higher predicted counts (the coast on the Norwegian Sea and Trondheimsfjorden), model 2 produced predictions with higher uncertainty (dark blue in the central panel), while on the mountains the uncer-
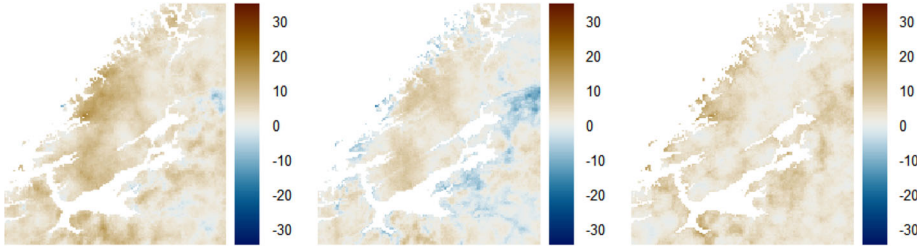
Figure 7. Differences in the posterior standard error of $Y_1(\mathbf{s}) + Y_2(\mathbf{s})$ for: model 1 and model 2 (left), model 3 and model 2 (center) and model 1 and model 3 (right).

tainty produced by model 3 was larger (light brown in the central panel). As a way to better appreciate the numerical differences between models, we explored the total predicted counts at the 113 sampling sites in Norway by comparing them against the observed counts (Fig. 8).

Figure 8 shows the comparison between the predicted and the observed values of total abundance of birds ($Y_1(\mathbf{s}) + Y_2(\mathbf{s})$). Model 1 and model 3 predict very similar values, and thus, we also compared the observed and predicted values of the counts gathered via point counts $Y_1(\mathbf{s})$ and line transects $Y_2(\mathbf{s})$ separately. Although model 1 and model 3 produce very similar predictions of total abundance of birds, model 3 predicted Y_1 and Y_2 separately more accurately. This is due to the inclusion of the GRF $\omega_2(\mathbf{s})$, as it makes it possible to better distribute the abundance between likelihoods and is flexible enough to capture more complex relationships between the census processes. We have highlighted the predicted and observed counts of the site located in the island of Linesøya (in red in Fig. 8) as this is a site where big discrepancies are observed between all the models. Model 1 and model 2 are not
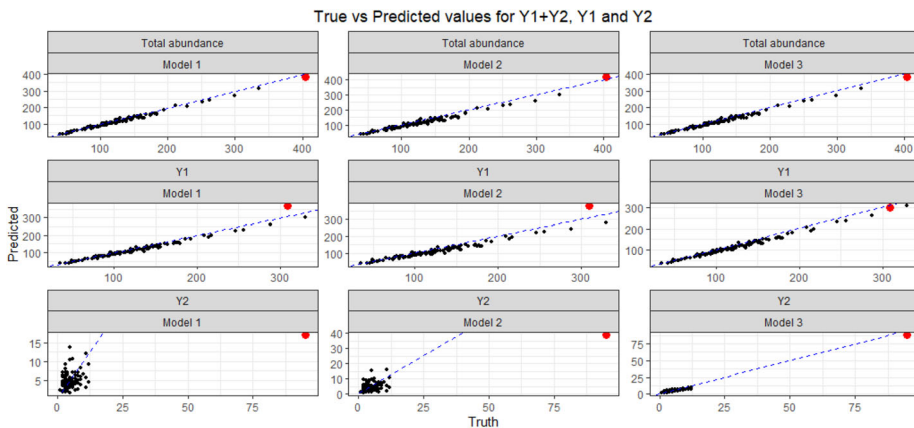
Figure 8. Comparison of observed vs predicted counts for: total abundance($Y_1(\mathbf{s}) + Y_2(\mathbf{s})$; top row), counts produced via point counts $Y_1(\mathbf{s})$ (middle row) and counts produced via line transect counts $Y_2(\mathbf{s})$ (bottom row). The performance of model 1, model 2 and model 3 is displayed in the first, second and third columns, respectively. A particular site with high total abundance of birds due to presence of gregarious species (in this case) that is only captured by one (the line transect in Norway) of the four census protocols is highlighted in red to allow for a quick assessment of discrepancies between the three models.

able to accurately predict the counts reported in this site by the line transect survey in Norway. This site is a special location where gregarious geese belonging to several species aggregate and form large gaggles (similar examples elsewhere might be sites with (multi-species) colonies, roosting sites or wetlands hosting thousands of waterbirds). Such information is only available if data from several census protocols are combined and properly analyzed— our new modeling framework can account for these differences, as our model 3 does in comparison with model 1 that assumes a linear relation.

Figure 9 shows the posterior mean of $\omega_1(\mathbf{s})$ for the three models, as well as $\omega_2(\mathbf{s})$ for model 3. $\omega_1(\mathbf{s})$ is, in general, similar for the three models. The largest difference occurs in $\omega_1(\mathbf{s})$ for model 2, which has a shorter spatial range in comparison to the other two models. In addition, the highest contribution of $\omega_2(\mathbf{s})$ occurs in Linesøya, an island where high total abundance of birds can be recorded during the line transects, due to high concentrations of geese from several species (see above). Such species form large groups of individuals (so called, gaggles) in some of the islands along the Norwegian coast. Lastly, we compared our three models in terms of goodness of fit and predictive performance (Table 4) using the measures of performance introduced in Sect. 3.4 and out-of-sample predictive performance measures such as RMSE after brute-force Leave-One-Out Cross Validation (CV), (Vehtari et al. 2016) and Leave-One-Site-Out CV. In the former CV scheme, we removed one data point at a time, while for the other we removed both the point and line transect counts. This procedures were computationally demanding, but feasible for our problem as it took 1.76 hours for model 1, 4.2 hours for model 2 and 4.1 hours for model 3.
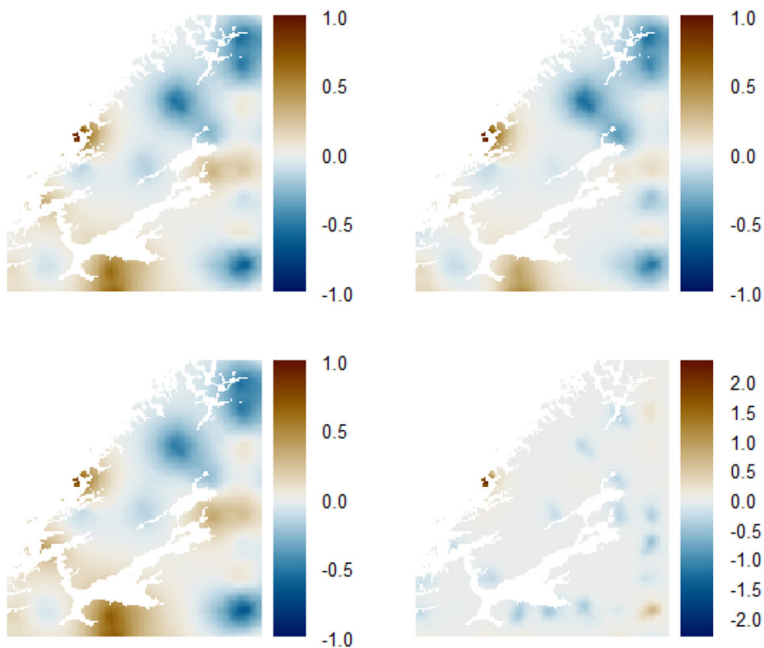

Figure 9.   Posterior mean of $\omega_1(\mathbf{s})$ for model 1 (upper left), model 2 (upper right) and model 3 (bottom left). Posterior mean of $\omega_2(\mathbf{s})$ for model 3 (bottom right).

Table 4. Measures of performance (see Sect. 3.3) for models 1, 2 and 3

Measure of performance	Model 1	Model 2	Model 3
DIC	2728.79	2751.04	**2603.82**
WAIC	2876.19	2876.96	**2593.69**
RMSE	165.60	145.00	**38.63**
LMPL	− 1673.50	− 1656.35	**− 1425.04**
Mean CRPS	27.21	25.93	**20.95**
RMSE (Leave-one-site-out CV)	**45.70**	46.39	45.88

In bold, the model with the best performance

The results show a considerable improvement in the goodness of fit when a second GRF to account for the particular characteristics in one of the observed data sources (line transects in Norway) is added. Moreover, the improvement in predictive performance of model 3 is exemplified by its low values of RMSE for the point count surveys in Norway, its high value of LMPL and its low CRPS for the point transect counts in Norway. The result of the leave-on-site-out CV shows small differences, but model 1 outperformed the other two models.

6. DISCUSSION AND CONCLUSIONS

The main goal of this paper was to introduce a modeling framework that allows us to model jointly multiple sources of information (count data) that are collected under different sampling protocols. We also presented a simple case study where we used this new methodology to estimate the total abundance of birds in mid-Scandinavia using bird counts in Norway and Sweden. These two countries have well-established bird monitoring programs, but differ in the sampling protocols. Therefore, we proposed a set of models that assumed the same coefficients for the fixed effects in each likelihood and a common GRF. The only difference between the different likelihoods is random intercepts in the linear predictor that aim at accounting for differences in the sampling protocols. For example, while the observed point counts in Norway have pairs of birds as the unit reported, Sweden reports individuals. Having different random intercepts makes possible to establish a proportional relation between the observed counts in the data sources. This is arguably a sensible choice since the biological processes that determine the abundance of species do not generally depend on national borders. Although the assumption of linear relation is reasonable for this case, it is also true that when working with real data allowing for some flexibility with respect to this assumption may correspond better to reality in most cases. This is why, we proposed a model that has a common GRF, but with a coefficient that explains how far we are from a linear relation. As seen in the exploratory analysis (Sect. 2), one of our data sources did not seem to follow the assumption of linear association with the other likelihoods. Hence, we suggested the inclusion of a second GRF to account for the differences of this likelihood. The inclusion of the second GRF, $\omega_2(\mathbf{s})$, was especially useful in our case as we do not have variables at the spatial point level that explicitly inform on the differences of the line count surveys in

Norway with respect to the other likelihoods. Simmonds et al. (2020) show the benefits of including an extra GRF to account for sources of bias in the sampling process of Citizen Science data. We assessed the performance of the three models when the key assumptions in the specification of each of them were not met in two simulation studies. The results of these simulations showed that a flexible specification performed similarly to the model that assumed a linear relation (model 1) when the latter model was used to generate the data. On the other hand, when the linear assumption was not met by the data generating model, the gap in performance between models became more evident. This suggests that using the models with flexible specification is always advised, regardless of the nature of the data. The estimates of the parameters in model 1 (the model assuming a linear relation between the observed counts) were biased and more uncertain than the estimates of the same parameters in the other two models. When a more complex scenario was proposed, model 3 (the model with two GRFs) clearly outperformed model 1 and model 2 in every comparison criteria. From the two simulation studies, we can conclude that model 3 is more robust than the other two models to misspecification of the functional form of the model. The parameters that showed higher differences in terms of bias and mean RMSE in the simulation study were the hyperparameters ζ_j. This might be caused by caused by the fact that these parameters are the only ones that are not constrained to be the same for all the likelihoods, and therefore, they are more sensitive to misspecification. A biased estimate of these hyperparameters might have an impact on the predictions of our models (total abundance of birds, in our case study) as these coefficients can be used as weighting of different likelihoods when computing the total abundance. The data of the simulation studies were also used to show why integrating the four sources of information is better for predicting the total counts of birds in more than one country (See Section S.1.2. of the Supplementary Information). We compared the predictive performance of a set of models that include (i) only one of the four sources of information, (ii) two sources of information (from the same country to predict abundance in a given location within the corresponding country—e.g., points and lines from Norway to predict within Norway), and (iii) the four sources of information (points and lines from both countries) (see Table S.2.). The results show that if the goal of the study is to produce predictions in more than one country, then integrating sources of information from both countries is recommended. If the goal of the study is to only produce within-country predictions, then integrating information for more than one country would not provide any additional benefit as the models with two sources of information performs as well as the models with the four sampling protocols. When we applied this methodology to the case study of estimating total bird abundance in mid-Scandinavia, we found some very high counts on the island of Linesøya (compared to elsewhere in the region). This count was recorded during a line transect sampling, which model 1 and model 2 failed to explicitly account for. This is arguably why the differences in goodness of fit between model 1 and model 2 were negligible. The inclusion of a second GRF in model 3 to explain extra complexity (in this case, the line counts in Norway that may produce large number of birds) made sense for our research problem since it was able to explain the large counts in Linesøya, when a large number of geese congregate around these islands. Adding GRFs to the likelihoods in order to account for particularities of each observed response seemed useful and practical in other cases when researchers need to account for complexity that

can not be explained with available covariate information. However, this addition should have a clear justification and be applied with caution since giving an ecological interpretation to this random effect may not be a trivial task. Our modeling framework offers, thus, advantages to integrate data from surveys with different sampling protocols and disjoint spatial locations. In its most simple parametrization, it does not explicitly account for any factor that affects the observed total abundance (i.e., detection). For example, in our case study, we have assumed these factors are negligible. However, this modeling framework is flexible enough to explicitly account for factors that influence the observed abundance. As shown in Sect. 3, these factors can be accounted for by explaining each of the terms ζ_j in the models proposed as a function of fixed and random effects that affect the observation process. Given the complexity of the models, identifiability issues may arise if the parameters that explain the effect of the factors related to the observation process are not constrained. This issue can be overcome by integrating data that inform on these parameters, or informative prior knowledge about them. The proposed framework does not explicitly accommodate species-specific characteristics. In our case study, it was not necessary as we assumed all the species have the same weight on the estimated total abundance. However, this modeling framework can work for a broader range of goals. For example, if one or a group of species are of interest when studying anthropogenic impacts on birds (e.g., total raptor counts (De Lucas et al. 2008)), the raw data can be preprocessed according to the purpose of the study. If the goal is to model one species of concern, then getting the subset of the raw data that belong to this species would suffice to apply our methodology and obtain satisfactory results. If, in another case, the question we want to solve is linked to the risk of collision of birds with powerlines (e.g., D'Amico et al. 2019) or rotor blades in wind farms (see De Lucas et al. 2008), we can account for the differences in sensitivity between species (for example soaring raptors, which are proportionally scarce in common bird monitoring schemes, are more sensitive than other bird species). Thus, one would multiply (apply weights) the count of each species in the dataset by a 'species-specific sensitivity factor' to that particular human impact (in this case, counts of raptor species would have a larger weight than other species). Then, one would proceed by summing up the new weighted counts to obtain a 'total weighted abundance of birds' at each census site. Our methodology, thus, can provide estimates of such a total weighted abundance across the entire region of interest and maps of 'sensitivity-adjusted hotspots.' An open question would be then, how to decide the values of these weights, which might be decided based on, for example, expert opinion, traits databases (Tobias et al. 2022) and published literature (D'Amico et al. 2019). A limitation of this modeling framework is that it lies in the category of purely spatial SDMs and thus it is not possible to explicitly account for any potential temporal variation at small (e.g., within a day) or large (e.g., across years) scale. In our case study, this was not a major concern as the temporal span of our data (14 years) is not considered a period in which the distribution of the total abundance of birds has varied a lot in the study region. The ultimate goal of developing this methodology is to integrate the different sources of bird count data to predict total abundance of birds across Norway, information that will be used in further studies of human impact on biodiversity, including predicting bird mortality hotspots due to powerlines and wind farms (Bernardino et al. 2018; Bevanger 1995, 2001; Serrano et al. 2020). Therefore, achieving a good predictive performance of our models is

of paramount importance to properly assess the vulnerability of different regions to human development based on the total local abundance of birds. Although we found differences in goodness of fit between the three models, the differences in predictive performance were small. However, a flexible model specification seemed the best choice for ensuring good predictions. For example, model 3 (which included $\omega_2(\mathbf{s})$ to account for particularities of the line counts in Norway) yields the most accurate predictions at the observed locations in Norway. This is associated with the extra complexity found between line transects and point counts in Norway, which unlike the two sampling protocols in Sweden did not have a clear linear relation, as they are only complementary to one another. In conclusion, in this paper we propose models to integrate multiple professional surveys with differences in their sampling protocols. These differences are usually determined by the country of origin of the data (sampling protocol) or by the specific targets of each monitoring scheme. The INLA-SPDE approach implemented in the R-INLA package makes it straightforward to perform full Bayesian inference for models that integrate multiple sources of information, even if they are not standardized or report the observed counts in different units. A natural extension of this work is the application of the proposed modeling framework to solve a broader range of ecological questions at larger geographical scales or for species with poor data (**Buckland and Johnston**) that incorporate more sources of information given its convenience and simple implementation.

REFERENCES

Araújo MB, Anderson RP, Márcia Barbosa A, Beale CM, Dormann CF, Early R, Garcia RA, Guisan A, Maiorano L, Naimi B, O'Hara RB, Zimmermann NE, and Rahbek C (2019). Standards for distribution models in biodiversity assessments. Sci Adv, 5(1)

Banerjee S, Carlin BP, Gelfand AE (2015) Hierarchical modeling and analysis for spatial data. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. CRC Press, 2ed. edition

Bernardino J, Bevanger K, Barrientos R, Dwyer J, Marques A, Martins R, Shaw J, Silva J, Moreira F (2018) Bird collisions with power lines: state of the art and priority areas for research. Biol Cons 222:1–13

Bevanger K (1995) Estimates and population consequences of tetraonid mortality caused by collisions with high tension power lines in Norway. J Appl Ecol 32(4):745–753

Bevanger K, Bartzke G, Brøseth H, Dahl E, Gjershaug J, Hanssen F, Jacobsen K-O, Kleven O, Kvaløy P, May R, Meås R, Nygård T, Refsnæs S, Stokke S, Thomassen J (2014) Optimal design and routing of power lines: ecological, technical and economic perspectives (optipol). final report: findings 2009–2014

Bevanger K (2001) Bird collisions with power lines - an experiment with ptarmigan (lagopus spp.). Biol Cons 99(3):341–346

Blangiardo M, Cameletti M (2015) Spatial and spatio-temporal Bayesian models with R-INLA. Wiley, New York

Bradter U, Ozgul A, Griesser M, Layton-Matthews K, Eggers J, Singer A, Sandercock BK, Haverkamp PJ, Snäll T (2021) Habitat suitability models based on opportunistic citizen science data: Evaluating forecasts from alternative methods versus an individual-based model. Divers Distrib 27(12):2397–2411

Brodie SJ, Thorson JT, Carroll G, Hazen EL, Bograd S, Haltuch MA, Holsman KK, Kotwicki S, Samhouri JF, Willis-Norton E, Selden RL (2020) Trade-offs in covariate selection for species distribution models: a methodological comparison. Ecography 43(1):11–24

Buckland S, Johnston A (2017) Monitoring the biodiversity of regions: key principles and possible pitfalls. Biol Cons 214:23–34

Cameletti M, Gómez-Rubio V, Blangiardo M (2019) Bayesian modelling for spatially misaligned health and air pollution data through the INLA-SPDE approach. Spat Stat 31:100353

De Lucas M, Janss GFE, Whitfield DP, Ferrer M (2008) Collision fatality of raptors in wind farms does not depend on raptor abundance. J Appl Ecol 45(6):1695–1703

D'Amico M, Martins RC, Álvarez Martínez JM, Porto M, Barrientos R, Moreira F (2019) Bird collisions with power lines: prioritizing species and areas by estimating potential population-level impacts. Divers Distrib 25(6):975–982

Fuglstad G-A, Simpson D, Lindgren F, Rue H (2019) Constructing priors that penalize the complexity of gaussian random fields. J Am Stat Assoc 114(525):445–452

Gelman A, Hwang J, Vehtari A (2014) Understanding predictive information criteria for Bayesian models. Stat Comput 24(6):997–1016

Gneiting T, Raftery AE (2007) Strictly proper scoring rules, prediction, and estimation. J Am Stat Assoc 102(477):359–378

Gomez-Rubio V (2020) Bayesian inference with INLA. CRC Press, Boca Raton

Gruss A, Thorson JT (2019) Developing spatio-temporal models using multiple data types for evaluating population trends and habitat usage. ICES J Mar Sci 76(6):1748–1761

Howard C, Stephens PA, Pearce-Higgins JW, Gregory RD, Willis SG (2014) Improving species distribution models: the value of data on abundance. Methods Ecol Evol 5(6):506–513

Järvinen O, Väisänen RA (1983) Correction coefficients for line transect censuses of breeding birds. Ornis Fennica 60(4):97–104

Johnston A, Fink D, Reynolds MD, Hochachka WM, Sullivan BL, Bruns NE, Hallstein E, Merrifield MS, Matsumoto S, Kelling S (2015) Abundance models improve spatial and temporal prioritization of conservation resources. Ecol Appl 25(7):1749–1756

Kéry M, Royle JA (2009). Inference about species richness and community structure using species-specific occupancy models in the national swiss breeding bird survey MHB, pp 639–656. Springer US, Boston, MA

Krainski ET (2019) Advanced spatial modeling with stochastic partial differential equations using R and INLA. CRC Press, Boca Raton

Kålås J (2010). The 2010 norwegian red list for species

Kålås J, Husby M(2002). Ekstensiv overvaking av terrestre fugl i norge

Lehikoinen A, Brotons L, Calladine J, Escandell CT, Flousek V, Grueneberg J, Haas C, Harris F, Herrando S, Jiguet HM, Kålås F, Lindström JA, Lorrilliere A, Molina R, Pladevall B, Calvi C, Sattler G, Schmid T, Trautmann H (2019) Declining population trends of european mountain birds. Glob Change Biol 25(2):577–588

Lehtomäki J, Moilanen A (2013) Methods and workflow for spatial conservation prioritization using zonation. Environ Modell Softw 47:128–137

Lindgren F, Rue H, Lindström J (2011) An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. J R Stat Soc Ser B (Stat Methodol) 73(4):423–498

Lindström Å, Green M, Husby M, Kålås JA, Lehikoinen A, Stjernman M, et al. (2019). Population trends of waders on their boreal and arctic breeding grounds in northern europe. Wader Study

Lindstrom A, Green M, Paulson G, Smith HG, Devictor V (2013) Rapid changes in bird community composition at multiple temporal and spatial scales in response to recent climate change. Ecography 36(3):313–322

Lissovsky AA, Dudov SV, Obolenskaya EV (2021) Species-distribution modeling: advantages and limitations of its application. 1. General approaches. Biol Bull Rev 11(3):254–264

Massimino D, Johnston A, Gillings S, Jiguet F, Pearce-Higgins JW (2017) Projected reductions in climatic suitability for vulnerable British birds. Clim Change 145(1):117–130

May R, Middel H, Stokke BG, Jackson C, Verones F (2020) Global life-cycle impacts of onshore wind-power plants on bird richness. Environ Sustain Indicators 8:100080

Miller DAW, Pacifici K, Sanderlin JS, Reich BJ (2019) The recent past and promising future for data integration methods to estimate species' distributions. Methods Ecol Evol 10(1):22–37

Pettit LI (1990) The conditional predictive ordinate for the normal distribution. J R Stat Soc Ser B (Methodol) 52(1):175–184

Roksvåg T, Steinsland I, Engeland K (2020). A geostatistical two field model that combines point observations and nested areal observations, and quantifies long-term spatial variability–a case study of annual runoff predictions in the voss area

Rue H, Held L (2005). Gaussian Markov random fields: theory and applications. Monographs on statistics and applied probability 104. Chapman & Hall CRC, 1 edition

Rue H, Martino S, Chopin N (2009) Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. J R Stat Soc Ser B (Stat Methodol) 71(2):319–392

Selle ML, Steinsland I, Hickey JM, Gorjanc G (2019) Flexible modelling of spatial variation in agricultural field trials with the r package Inla. Theor Appl Genet 132(12):3277–3293

Selle ML, Steinsland I, Powell O, Hickey JM, Gorjanc G (2020) Spatial modelling improves genetic evaluation in smallholder breeding programs. Genet Sel Evol 52(1):1–17

Serrano D, Margalida A, Pérez-García JM, Juste J, Traba J, Valera F, Carrete M, Aihartza J, Real J, Mañosa S, Flaquer C, Garin I, Morales MB, Alcalde JT, Arroyo B, Sánchez-Zapata JA, Blanco G, Negro JJ, Tella JL, Ibañez C, Tellería JL, Hiraldo F, Donázar JA (2020) Renewables in Spain threaten biodiversity. Science 370(6522):1282–1283

Simmonds EG, Jarvis SG, Henrys PA, Isaac NJB, O'Hara RB (2020) Is more data always better? A simulation study of benefits and limitations of integrated distribution models. Ecography 43(10):1413–1422

Simpson D, Illian JB, Lindgren F, Sørbye SH, Rue H (2016) Going off grid: computationally efficient inference for log-Gaussian Cox processes. Biometrika 103(1):49–70

Soultan A, Pavón-Jordán D, Bradter U, Sandercock BK, Hochachka WM, Johnston A, Brommer J, Gaget E, Keller V, Knaus P, Aghababyan K, Maxhuni Q, Vintchevski A, Nagy K, Raudonikis L, Balmer D, Noble D, Leitão D, Øien IJ, Shimmings P, Sultanov E, Caffrey B, Boyla K, Radišić D, Lindström Å, Velevski M, Pladevall

C, Brotons L, Karel Š, Rajković DZ, Chodkiewicz T, Wilk T, Szép T, van Turnhout C, Foppen R, Burfield I, Vikstrøm T, Mazal VD, Eaton M, Vorisek P, Lehikoinen A, Herrando S, Kuzmenko T, Bauer H-G, Kalyakin MV, Voltzit OV, Sjeničić J, Pärt T (2022) The future distribution of wetland birds breeding in Europe validated against observed changes in distribution. Environ Res Lett 17(2):024025

Soykan CU, Sauer J, Schuetz JG, LeBaron GS, Dale K, Langham GM (2016) Population trends for north American winter birds based on hierarchical models. Ecosphere 7(5):e01351

Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A (2002) Bayesian measures of model complexity and fit. J R Stat Soc Ser B (Stat Methodol) 64(4):583–639

Tobias JA, Sheard C, Pigot AL, Devenish AJM, Yang J, Sayol F, Neate-Clegg MHC, Alioravainen N, Weeks TL, Barber RA, Walkden PA, MacGregor HEA, Jones SEI, Vincent C, Phillips AG, Marples NM, Montaño-Centellas FA, Leandro-Silva V, Claramunt S, Darski B, Freeman BG, Bregman TP, Cooney CR, Hughes EC, Capp EJR, Varley ZK, Friedman NR, Korntheuer H, Corrales-Vargas A, Trisos CH, Weeks BC, Hanz DM, Töpfer T, Bravo GA, Remeš V, Nowak L, Carneiro LS, Moncada R (2022) Avonet: morphological, ecological and geographical data for all birds. Ecol Lett 25(3):581–597

Vehtari A, Mononen T, Tolvanen V, Sivula T, Winther O (2016) Bayesian leave-one-out cross-validation approximations for gaussian latent variable models. J Mach Learn Res 17:103:1-103:38

Watanabe S (2010) Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. J Mach Learn Res 11:3571–3594

Yu H, Cooper AR, Infante DM (2020) Improving species distribution model predictive accuracy using species abundance: application with boosted regression trees. Ecol Modell 432:109202